 doi 10.3969/j.issn.1673-4254.2015.11.24　　J South Med Univ, 2015, 35(11): 1638-1642

技术方法

# 贝叶斯期中分析与经典期中分析在成组序贯设计中的比较

原玲玲[1],詹志颖[2],谭旭辉[2]
南方医科大学[1]人文与管理学院,[2]生物统计学系,广东 广州 510515

**摘要:目的** 比较贝叶斯期中分析与经典方法的期中分析的差异。**方法** 以对照组(Control)和试验组(Treatment)的两样本均数比较为分析目的,即$\theta=\mu_T-\mu_C$($\theta$越大疗效越好),建立$H_0$:$\theta\leqslant 0$;$H_1$:$\theta>0$的优效性假设检验(拒绝$H_0$,即支持处理组疗效)。按照成组序贯设计的数据要求,在每次期中分析时刻,计算各种先验分布的贝叶斯期中分析Ⅰ类错误、功效、平均样本量、平均阶段数等指标。**结果** 在Pocock和O'Brien & Fleming设计中,Skeptical先验和Handicap先验的Ⅰ类错误$\varepsilon$均能控制在0.05左右。当O' Brien & Fleming和Pocock方法功效在80%时,基于Handicap先验和Skeptical先验的贝叶斯功效相对来说明显较低,而基于Non-informative先验和Enthusiastic先验的贝叶斯功效则明显较高。**结论** Skeptical先验和Handicap先验的贝叶斯期中分析能较好的控制Ⅰ类错误$\varepsilon$,基于Skeptical先验和Handicap先验的贝叶斯期中分析相对于O'Brien & Fleming方法均能够明显增加试验提前终止的可能性,而对于Pocock方法则没有太大实际意义。
**关键词:**成组序贯设计;贝叶斯;期中分析;先验分布

# Comparison of Bayesian interim analysis and classical interim analysis in group sequential design

*YUAN Lingling[1], ZHAN Zhiying[2], TAN Xuhui[2]*
*[1]School of Humanities and Management, [2]Department of Biostatistics, Southern Medical University, Guangzhou 510515, China*

**Abstract: Objective** To explore the differences between the Bayesian interim analysis and the classical interim analysis. **Methods** To compare the means of two independent samples between control and treatment, superior hypothesis test was established. In line with the data requirements for group sequential design, Type I error of Bayesian interim analysis based on various prior distributions, Power, Average Sample Size and Average Stage were estimated in the interim analysis. **Results** In the Pocock and O' Brien & Fleming designs, the Type I errors in the Bayesian interim analysis based on the skeptical prior distribution and the handicap prior distribution were controlled at around 0.05. When the powers of these two classical designs were both 80%, Bayesian powers of the skeptical prior distribution and the handicap prior distribution were markedly lower. The powers of the non-informative prior distribution and the enthusiastic prior distribution were distinctly higher than 80%. **Conclusion** In the Bayesian interim analysis based on the skeptical prior distribution and the handicap Prior distribution, the Type I errors can be well controlled. Bayesian interim analyses using these two prior distributions, compared with the analysis adopting the O' Brien & Fleming method, can markedly increase the possibility of ending the clinical trials ahead of time. The Bayesian interim analyses based on these two distributions do not have practical value for group sequential design of the Pocock method.
**Key words:** group sequential design; Bayes; interim analysis; prior distribution

在临床试验领域,以往试验的先验信息没有被充分利用是一个普遍现象,而贝叶斯统计[1-2]是一种能充分利用先验信息的统计方法。在药物临床试验中,期中分析通常能够有效的降低试验所需的样本量并提前完成试验,从而为药物的上市节省大量的时间和资源。而结合了先验信息的贝叶斯期中分析在临床试验中相对于传统期中分析是否可行可靠、能否进一步提高效率等相关疑问亟待明确。

**收稿日期:**2015-11-01
**基金项目:**国家自然科学基金(81302515);广东省自然科学基金(S2013040013902);高等学校博士学科点专项科研基金(20134433120023)
Supported by National Natural Science Foundation of China (81302515).
**作者简介:**原玲玲,博士,助理研究员,E-mail: txhyll@163.com
**通信作者:**谭旭辉,博士,副教授,E-mail: txh8287@sina.com

因此,为了了解贝叶斯期中分析与成组序贯设计[3-4]的经典期中分析[5]的差异,揭示两种方法在具体应用上的优劣势,定量评估贝叶斯结果的可靠性,本文将对贝叶斯期中分析与经典期中分析进行比较,重点对几种常见的先验分布的影响进行评估,为贝叶斯期中分析在临床试验领域的应用提供理论参考,也为今后药物临床试验的期中分析提供新的思路。

## 1 方法与原理

### 1.1 成组序贯设计

成组序贯设计是一种在试验正式结束前,能提供多次期中分析的设计。成组序贯方法是由Pocock于1977年提出,在此基础上Kittelson和Emerson于1999年提

出了Unified family方法[6]，其包括Pocock、O'Brien-Fleming、Power family等方法。其中Power family方法是由Wang和Tsiatis于1987提出[7]，其通过参数Power值的改变来得到不同的临界值。从设计方案角度来说，成组序贯设计是将整个试验划分成$K$个连续的时间段，每个时间段内每个处理组都有$n$个受试者加入。当第$k$个阶段($k$=1,2,…,$K$)结束后，把之前各个阶段试验结果累积起来进行一次统计分析。对于拒绝性的成组序贯设计，如果拒绝$H_0$则试验结束，否则继续下一阶段试验。当其处于最后一个试验阶段时，其结果不是接受$H_0$，就是拒绝$H_0$。

1.2 贝叶斯统计

假设参数$\theta$，其概率分布为$\pi(\theta)$，观察值$y_1,\ldots,y_n$来自概率密度函数为$P(y|\theta)$，则：

$$p(\theta|y) = p(\theta,y)/p(y) = p(y|\theta)\pi(\theta)/\int p(y|\theta)\pi(\theta)\mathrm{d}\theta$$

$$\text{设 } L(\theta)\propto p(y|\theta)\text{，则 } p(\theta|y)\propto L(\theta)\pi(\theta)$$

由以上过程也可发现，贝叶斯统计推断均是基于后验概率分布$p(\theta|y)$。而贝叶斯统计与经典统计的主要区别就在于先验信息的使用。在贝叶斯统计中，先验一般都具有主观性，然而为了减少主观性对后验概率的影响，也提出了一些较"客观"的先验分布[8]，例如广义先验分布、无信息先验分布、共轭先验分布以及有信息先验分布等。在一些较复杂的具体分析中，贝叶斯统计需要应用MCMC方法模拟产生基于后验分布的样本，并以此来估计特征统计量。

1.3 贝叶斯期中分析

贝叶斯期中分析是基于试验过程中所得到的后验概率分布，并在此基础上作出统计推断。在该过程中不需要事先指定终止标准以及所需样本量，仅仅应用后验$P$值来进行期中监测[9]。而经典的期中分析是建立在"Neyman-Pearson"理论上，其重点是在于控制Ⅰ和Ⅱ类错误。

ICH-GCP将期中分析定义为：在一个临床试验正式完成之前的任何时间内，为了比较组间的有效性或安全性而进行的分析。依据后验概率$P$的期中分析，如果小于设定的临界值$\varepsilon$，即$p(H_0|data)\leqslant\varepsilon$(为了与经典统计Ⅰ类错误$\alpha$区分，记为$\varepsilon$)，则拒绝$H_0$。因此，先验分布的选择对于贝叶斯期中分析是至关重要的。为了让试验支持者或试验怀疑者能够达成较一致的结论，先验分布的选择有如下两个原则[10]：(1)如果是阳性结局(positive result)的试验，其$P$值的计算应基于Skeptical先验；(2)如果是阴性结局(negative result)的试验，其$P$值的计算应基于Enthusiastic先验。

在以上两原则的条件下，仍以较大的概率支持相应的结局，则可下支持相应结局的结论。

1.4 期中分析之间的比较设计

在临床试验中，Pocock和O' Brien & Fleming[11]期中分析方法是成组序贯设计中比较常见的两种方法。在对比研究中，我们以经典成组序贯为设计框架，按照频率学派的观点，在功效，Ⅰ类错误以及实际所需样本量等方面，将贝叶斯期中分析与常见的两种经典期中分析方法进行比较。由于成组序贯设计在实施前需确定一些参数，如$\alpha$、$\beta$、$\delta$、以及期中分析次数等，我们事先定义期中分析次数为$k$，总分析次数为$K$(即$K$=$k$+1)。以对照组和试验组的两样本均数比较为分析目的，即$\theta=\mu_T-\mu_C$($\theta$越大疗效越好)，建立$H_0$:$\theta\leqslant0$；$H_1$:$\theta>0$的优效性假设检验。对于两组的试验模拟数据，可由样本量公式计算出第$k$阶段每组所需样本量(假设两组样本量相等)。我们以$N(\mu_T,\sigma_T^2)$、$N(\mu_C,\sigma_C^2)$表示试验和对照组总体，从中随机模拟生成各个阶段所需的样本试验数据。为了简化设计，我们假设$\sigma_T=\sigma_C$，$\mu_T>\mu_C$，即两处理组有共方差。在模拟比较中，较"客观"的先验分布选择Enthusiastic先验、Skeptical先验、Handicap先验和Non-informative先验(无信息先验)。Pocock和O'Brien & Fleming两种方法的参数设定如表1。

基于经典期中分析结果，提出了评价贝叶斯期中分析效果的两个指标，即贝叶斯阳性符合率和贝叶斯阴性符合率[12]。贝叶斯阳性符合率计算公式如公式(1)：

$$C_p=\frac{N_p^{CB}}{N_p^C} \tag{1}$$

假设在有差异的两总体中，进行了$N_p$次成组序贯模拟；$N_p^C$:在$N_p$试验中，经典期中分析$N_p^C$次拒绝$H_0$；$N_p^B$:在$N_p$试验中，贝叶斯期中分析$N_p^B$次拒绝$H_0$；$N_p^{CB}$:在经典期中分析$N_p^C$次拒绝$H_0$的试验中，贝叶斯期中分析也拒绝$H_0$的次数。

从公式(1)，我们可推出，经典期中分析的功效$Power_C=N_p^C/N_p$；贝叶斯期中分析的功效$Power_B=N_p^B/N_p$

贝叶斯阴性符合率计算公式如公式(2)：

$$C_n=\frac{N_n^{CB}}{N_n^C} \tag{2}$$

假设在无差异的两总体中，进行了$N_n$次成组序贯模拟；$N_n^C$:在$N_n$试验中，经典期中分析$N_n^C$次不拒绝$H_0$；$N_n^B$:在$N_n$试验中，贝叶斯期中分析$N_n^B$次不拒绝$H_0$；$N_n^{CB}$:在经典期中分析$N_n^C$次不拒绝$H_0$的试验中，贝叶斯期中分析也不拒绝$H_0$的次数。

从公式(2)，我们可推出，经典期中分析的Ⅰ类错误$\alpha=1-N_n^C/N_n$；按照频率学派观点，贝叶斯期中分析的Ⅰ类错误$\varepsilon=1-N_n^B/N_n$。与此同时，我们也可计算另外

两个指标，即贝叶斯期中分析的平均样本量(Average Sample Size)记为 $ASS_p$ 和平均终止阶段数(Average Stage)，记为$AS_p$。

表1 参数设定

Tab.1 Parameter setting

| Parameter | Value($\beta$=0.2) | |
|---|---|---|
| $\alpha$ | 0.05 | 0.05 |
| $\sigma_T=\sigma_C=\sigma$ | 10 | 18 |
| $\theta=\mu_T-\mu_C(\delta)$ | 2 | 3 |

## 2 结果

### 2.1 两总体均数相同

在两总体均数相同条件下，我们将成组序贯设计按以下参数进行设定(表2)。

在O_Set5_1和O_Set5_2(即O' Brien & Fleming方法)参数设置下，不同先验分布和贝叶斯期中分析阴性符合率以及Ⅰ类错误如表3。在P_Set5_1和P_Set5_2(即Pocock方法)参数设置下，不同先验分布和贝叶斯期中分析阴性符合率以及Ⅰ类错误如表4。

由以上结果，我们可以看出在Enthusiastic先验和Non-informative先验的条件下，贝叶斯期中分析的I类错误 均明显大于0.05；其中在Non-informative先验条件下，5阶段 在0.10以上。在Handicap先验和Skeptical先验条件下，Ⅰ类错误 均能控制在0.05左右。当总阶段数调整为3或者2时，贝叶斯期中分析阴性符合率和Ⅰ类错误与5阶段条件下的结果相似。

表2 O'Brien & Fleming和Pocock方法参数设定(两总体均数相等)

Tab.2 Parameter setting in the O'Brien & Fleming method and the Pocock method (two equal population means)

| | Condition | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Iteration | $\beta$ | $\sigma$ | Stage(K) | Stopping | $\mu_T$ | $\mu_C$ | $\delta$ |
| O_Set5_1 | 300 | 0.20 | 10 | 5 | Reject | 6 | 6 | 2 |
| O_Set5_2 | 300 | 0.20 | 18 | 5 | Reject | 6 | 6 | 3 |
| P_Set5_1 | 300 | 0.20 | 10 | 5 | Reject | 6 | 6 | 2 |
| P_Set5_2 | 300 | 0.20 | 18 | 5 | Reject | 6 | 6 | 3 |

表3 O'Brien & Fleming方法下的贝叶斯阴性符合率、Ⅰ类错误

Tab.3 Bayesian negative coincidence rate and Type I error in the O'Brien & Fleming method

| Method | O_Set5_1 | | O_Set5_2 | |
|---|---|---|---|---|
| | $C_n$(%) | $\alpha/\varepsilon$(%) | $C_n$(%) | $\alpha/\varepsilon$(%) |
| Non-informative Prior | 92.01 | 11.67 | 87.76 | 16.33 |
| Skeptical Prior | 99.65 | 3.33 | 99.30 | 4.00 |
| Enthusiastic Prior | 55.21 | 47.00 | 57.69 | 45.00 |
| Handicap Prior | 99.65 | 3.00 | 99.30 | 3.33 |
| O'Brien & Fleming | - | 4.00 | - | 4.67 |

表4 Pocock方法下的贝叶斯阴性符合率、Ⅰ类错误

Tab.4 Bayesian negative coincidence rate and Type I error in the Pocock method

| Method | P_Set5_1 | | P_Set5_2 | |
|---|---|---|---|---|
| | $C_n$(%) | $\alpha/\varepsilon$(%) | $C_n$(%) | $\alpha/\varepsilon$(%) |
| Non-informative Prior | 93.15 | 9.33 | 89.31 | 13.67 |
| Skeptical Prior | 99.32 | 2.67 | 98.62 | 3.00 |
| Enthusiastic Prior | 61.99 | 39.67 | 57.93 | 44.00 |
| Handicap Prior | 99.32 | 2.67 | 98.97 | 2.67 |
| Pocock | - | 2.67 | - | 3.33 |

### 2.2 两总体均数不同

在两总体均数不同条件下，我们将成组序贯设计按以下参数进行设定(表5)。

在O_Set5_D1、O_Set5_D2、P_Set5_D1和P_Set5_D2参数设置下，几种常见先验分布的贝叶斯期中分析平均阶段数、平均样本量、功效和阳性符合率(表6,7)。

通过以上5阶段的结果可以看出，当O'Brien & Fleming功效在80%左右时，基于Handicap先验和Skeptical先验的贝叶斯功效相对较低，而基于Enthusiastic先验和Non-informative先验的贝叶斯功效明显较高。此外，几种先验分布的贝叶斯期中分析所需的平均阶段数和平均样本量均比O'Brien & Fleming方法要低，而基于Handicap先验和Skeptical先验的贝叶斯期中分析均能提前终止试验。当Pocock功效在80%时，Skeptical和Handicap两种先验分布的贝叶斯功效仍旧明显较低，而Enthusiastic先验和Non-informative先验的贝叶斯功效均比Pocock功效要高。与此同时，基于Handicap先验和Skeptical先验的贝叶斯期中分析相对于Pocock方法而言并不能明显的提前终止试验。

## 3 讨论

### 3.1 两总体均数相同情况

从实际模拟效果来看，在O'Brien & Fleming和Pocock期中分析方法下，使用Enthusiastic先验和Non-

表5 O'Brien & Fleming和Pocock方法参数设定(两总体均数不等)

Tab.5 Parameter setting in the O' Brien & Fleming method and the Pocock method (two unequal population means)

| | Condition | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Iteration | $\beta$ | $\sigma$ | Stage(K) | Stopping | $\mu_T$ | $\mu_C$ | $\delta$ |
| O_Set5_D1 | 300 | 0.20 | 10 | 5 | Reject | 8 | 6 | 2 |
| O_Set5_D2 | 300 | 0.20 | 18 | 5 | Reject | 9 | 6 | 3 |
| P_Set5_D1 | 300 | 0.20 | 10 | 5 | Reject | 8 | 6 | 2 |
| P_Set5_D2 | 300 | 0.20 | 18 | 5 | Reject | 9 | 6 | 3 |

表6 O'Brien & Fleming方法下的贝叶斯平均样本量、平均阶段数、阳性符合率和功效

Tab.6 Average sample size, average stage, Bayesian positive coincidence rate and power in the O' Brien & Fleming method

| Method | O_Set5_D1 | | | | O_Set5_D2 | | | |
|---|---|---|---|---|---|---|---|---|
| | $ASS_p$ | $AS_p$ | $C_p$(%) | Power (%) | $ASS_p$ | $AS_p$ | $C_p$(%) | Power (%) |
| Non-informative Prior | 352.53 | 2.73 | 100.00 | 85.00 | 489.09 | 2.64 | 100.00 | 86.33 |
| Skeptical Prior | 469.91 | 3.65 | 89.18 | 69.67 | 663.69 | 3.58 | 92.05 | 73.67 |
| Enthusiastic Prior | 187.30 | 1.45 | 100.00 | 97.67 | 295.93 | 1.59 | 100.00 | 96.00 |
| Handicap Prior | 474.19 | 3.68 | 87.88 | 68.00 | 666.16 | 3.59 | 89.54 | 71.67 |
| O'Brien & Fleming | 498.21 | 3.87 | - | 77.00 | 703.74 | 3.79 | - | 79.67 |

表7 Pocock方法下的贝叶斯平均样本量、平均阶段数、阳性符合率和功效

Tab.7 Average sample size, average stage, Bayesian positive coincidence rate and power in the Pocock method

| Method | P_Set5_D1 | | | | P_Set5_D2 | | | |
|---|---|---|---|---|---|---|---|---|
| | $ASS_p$ | $AS_p$ | $C_p$(%) | Power (%) | $ASS_p$ | $AS_p$ | $C_p$(%) | Power (%) |
| Non-informative Prior | 388.85 | 2.50 | 100.00 | 91.33 | 547.25 | 2.45 | 100.00 | 92.00 |
| Skeptical Prior | 472.02 | 3.04 | 78.15 | 66.33 | 672.29 | 3.01 | 77.96 | 66.67 |
| Enthusiastic Prior | 227.95 | 1.46 | 100.00 | 97.67 | 297.02 | 1.33 | 100.00 | 99.33 |
| Handicap Prior | 473.05 | 3.04 | 77.31 | 65.33 | 675.28 | 3.02 | 75.92 | 64.33 |
| Pocock | 478.21 | 3.08 | - | 79.33 | 681.25 | 3.05 | - | 81.67 |

informative先验进行贝叶斯期中分析会增加I类错误的风险，而Handicap先验和Skeptical先验均可较好的控制Ⅰ类错误，因此这两种先验是比较合适的先验选择。Skeptical先验、Handicap先验和Non-informative先验均比Enthusiastic先验显著的增加贝叶斯阴性符合率，这意味着在原假设成立的条件下，基于Enthusiastic先验的贝叶斯期中分析结果很可能会和其他3种先验所得的结论不一致。这种现象提示我们，在保证试验因素能够继续进行研究的前提下，如果贝叶斯期中分析在Non-informative先验和Enthusiastic先验之间出现不一致的结论，则该研究可能是一个阴性研究[12]。我们还发现基于Handicap先验和Skeptical先验的贝叶斯阴性符合率均接近100%，从先验分布选择的角度来说，Skeptical先验是在阳性研究(Positive study)中建议使用的分布，这意味着其更倾向于“保守”，而Handicap先验的贝叶斯阴性符合率和Ⅰ类错误 均与Skeptical先验相近，因此可认为这两种先验的“保守”程度是比较接近的。此外，在大样本的条件下，不同先验分布的后验概率分布之间通常只有很微小的差异，因此大样本条件下基于各种先验的贝叶斯期中分析就等同于敏感度分析，其目的是评估试验数据支持结论的强度[12-13]。

3.2 两总体均数不同情况

从5阶段成组序贯设计所得的各种先验分布的功效结果来看，当O'Brien & Fleming和Pocock功效在80%时，基于Handicap先验和Skeptical先验的贝叶斯期中分析的功效相比较而言明显偏低；当O'Brien & Fleming功效逐渐增大时，以上两种先验分布的贝叶斯功效开始接近O'Brien & Fleming方法的功效。这是因为在参数相同条件下，如果先验方差较小，则基于该先验分布的贝叶斯期中分析更易得到较大的$P$值，这也意味着基于该先验的贝叶斯期中分析更加的保守，其功效也会相应的较低；如果两种先验方差的比值逐渐增大，

两者功效之间的差异则会逐渐增大。然而随着O'Brien & Fleming功效的逐渐提高，基于以上两种先验的贝叶斯功效也会逐渐接近O'Brien & Fleming的功效。这点也充分说明，在大样本条件下，不同先验分布下的后验概率分布之间通常只有微小的差异。此外，在O'Brien & Fleming和Pocock的功效较低的条件下，基于Handicap先验和Skeptical先验的贝叶斯阳性符合率均较低，这就意味着一方面基于这两种先验的贝叶斯期中分析的功效相对于Pocock和O'Brien & Fleming功效而言均较低；另一方面从频率学派角度来说，基于这两种先验的分析结果的可靠性也较低。值得一提的是，当O'Brien & Fleming功效在80%的时候，基于Skeptical先验和Non-informative先验的贝叶斯阳性符合率之间有较大差异。这种差异对于阳性研究来说不但意味着需要进行证实性的试验研究，而且也可能提示着试验因素可能没有被完全理解[12]。

对于O'Brien & Fleming方法来说，早期是很难拒绝零假设的，这也使得贝叶斯期中分析能提前终止试验的特点显得非常突出。然而在Ⅲ期临床试验中，为了对药物作有效性评价，成组序贯设计又通常会选择应用O'Brien & Fleming方法[14-15]，因此在验证有效性的临床试验中，当O'Brien & Fleming方法的功效越高，基于Handicap先验和Skeptical先验的贝叶斯期中分析的功效就会越接近O'Brien & Fleming方法的功效，在功效较高且相近的条件下，基于Handicap先验和Skeptical先验的贝叶斯期中分析均能增加试验提前结束的可能性。另外，基于Handicap先验和Skeptical先验的贝叶斯期中分析方法相对于Pocock方法而言并不能明显提前终止试验。这是因为5阶段Pocock法允许试验早期以较小的临界值(相当于早期设置了较大的名义检验水平，即名义$\alpha_1 = 0.0169$)来拒绝零假设，这也就意味着一些试验可能在早期(甚至在第1阶段)因为拒绝零假设而终止，正因为Pocock具有较早的拒绝$H_0$的特点，所以在提前终止试验方面贝叶斯期中分析相对于Pocock法而言并没有什么优势。

综上所述，当采用功效达到80%以上且早期不易拒绝零假设的传统成组序贯设计时，基于Skeptical先验和Handicap先验的贝叶斯期中分析既能使结果具有较强的可靠性，又能显著增加试验提前终止的可能性。

## 参考文献：


[1] Lin R, Yin G. Bayes factor and posterior probability: complementary statistical evidence to p-value[J]. Contemp Clin Trials, 2015, 44(1): 33-5.

[2] Freedman L. Bayesian statistical methods[J]. BMJ, 1996, 313(757): 569-70.

[3] Gao Z, Roy A, Tan M. Multistage adaptive biomarker-directed targeted design for randomized clinical trials[J]. Contemp Clin Trials, 2015, 42(1): 119-31.

[4] Lai TL, Liao OY, Kim DW. Group sequential designs for developing and testing biomarker-guided personalized therapies in comparative effectiveness research[J]. Contemp Clin Trials, 2013, 36(2): 651-63.

[5] Stephens RJ, Langley RE, Mulvenna P, et al. Interim results in clinical trials: do we need to keep all interim randomised clinical trial results confidential[J]. Lung Cancer, 2014, 85(2): 116-8.

[6] Kittelson JM, Emerson SS. A unifying family of group sequential test designs[J]. Biometrics, 1999, 55(3): 874-82.

[7] Wang SK, Tsiatis AA. Approximately optimal one-parameter boundaries for group sequential trials[J]. Biometrics, 1987, 43(1): 193-9.

[8] Berger J. The case for objective bayesian analysis[J]. Bayesian Anal, 2006, 1(3): 385-402.

[9] Daimon T. Predictive checking for Bayesian interim analyses in clinical trials[J]. Contemp Clin Trials, 2008, 29(5): 740-50.

[10] Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trial and health-care evaluation[M]. Chichester: John Wiley&Sons, 2004.

[11] Bartroff J, Lai TL, Shih MC. Sequential experimentation in clinical trials[M]. New York: Springer, 2013.

[12] Wijeysundera DN, Austin PC, Hux JE, et al. Bayesian statistical inference enhances the interpretation of contemporary randomized controlled trials[J]. J Clin Epidemiol, 2009, 62(1): 13-21.e5.

[13] Homs MY, Steyerberg EW, Eijkenboom WM, et al. Single-dose brachytherapy versus metal stent placement for the palliation of dysphagia from oesophageal Cancer: multicentre randomised trial[J]. Lancet, 2004, 364(9444): 1497-504.

[14] Skovlund E. Repeated significance tests on accumulating survival data[J]. J Clin Epidemiol, 1999, 52(11): 1083-8.

[15] Choi SC. Interim analyses and early termination of clinical trials[J]. J Biopharm Stat, 1997, 7(4): 533-43.


（编辑：孙昌朋）